Exhibit V

IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR 2004), Nov. 2-5, 2004, Washington, D.C.

FlightTracker: A Novel Optical/Inertial Tracker for Cockpit Enhanced Vision

Eric Foxlin, Yury Altshuler, Leonid Naimark and Mike Harrington
InterSense Inc.
(ericf/yurya/leonidn/mikeh@isense.com)

Abstract

One of the earliest fielded augmented reality applications was enhanced vision for pilots, in which a display projected on the pilot's visor provides geo-spatially registered information to help the pilot navigate, avoid obstacles, maintain situational awareness in reduced visibility, and interact with avionics instruments without looking down. This requires exceptionally robust and accurate head-tracking, for which there is not a sufficient solution yet available. In this paper, we apply miniature MEMS sensors to cockpit helmet-tracking for enhanced/synthetic vision by implementing algorithms for differential inertial tracking between helmet-mounted and aircraft-mounted inertial sensors, and novel optical drift correction techniques. By fusing low-rate inside-out and outside-in optical measurements with high-rate inertial data, we achieve millimeter position accuracy and milliradian angular accuracy, low-latency and high robustness using small and inexpensive sensors.

1. Introduction

1.1. Helmet-tracking requirements in cockpits

One of the key avionics technologies in a modern military aircraft is a helmet-tracking system (HTS). In most new programs being developed now, the HTS is tightly integrated with a helmet-mounted display (HMD) and the combined system is often abbreviated HMD/T. The HMD/T system can provide several crucial advantages to the pilot. Tracked helmet mounted displays can be used for enhanced or synthetic vision displays to aid pilot situational awareness, for human-machine interaction between the pilot and the aircraft systems, for heads-up presentation of critical flight data in head-stabilized or space-stabilized format, and for weapons cueing. In enhanced vision displays, computer graphic (CG) representations such as highway-in-the-sky, runway and building outlines, and obstacle highlighting symbols are overlaid directly on the pilot's view of the outside world, so very high angular accuracy must be achieved in the HMD/T system to portray a convincing registration. In synthetic vision, the head-tracker is used to slave externally mounted sensors (visible, IR, radar, etc) that are fused with one another and possibly also with a computer terrain model, to produce a visual display that replaces the pilot's natural vision. Therefore, a slightly lower accuracy might be acceptable as long as the tracker meets the resolution and responsiveness requirements of a virtual reality HMD system.

In weapons cueing, the pilot typically places a fixed targeting reticle at the center of the HMD over an enemy target, and then indicates it as a potential target. A line-of-sight (LOS) vector from the aircraft to the target is determined based on the HTS output and transferred to the high-off-boresight seeker head in the weapon, which begins tracking the target. As the seeker tracks the target, it continues to update the LOS vector to the aircraft computer, which combines the target LOS and the HTS data to present a targeting box in the HMD which appears to stay over the target as the pilot moves his head. For weapons cueing, the tracker errors in combination with all other aiming errors including those due to the HMD optics and refraction through the canopy must be kept to a fraction of a degree, resulting in specifications for helmet-tracker accuracy ranging from 5-7 mrad to as low as 1-3 mrad [1]. This extreme accuracy must be maintained with perfect reliability in the harsh environment of the moving vehicle, and a sufficient solution that does not require extensive calibration or pilot attention is still needed.

Of course, accuracy specifications alone cannot give a complete picture of the complexities involved in designing or selecting an HTS. Other important considerations include: sensor size and weight, ruggedness, reliability, robustness of the tracking data during harsh maneuvers or buffeting, graceful degradation for motions outside the motion box, ease of integration of the tracking components into the helmet and the cockpit, longevity, and ease and accuracy of calibration and boresighting procedures. Another tracker feature that needs to be considered is how many degrees-of-freedom (DOF) it needs to track. To establish a LOS vector to a target for weapons cueing, it is sufficient to measure the helmet azimuth and elevation, and some early helmet-mounted sights indeed lack roll or translation tracking. However, the curvature of the aircraft canopy can cause significant optical distortion that contributes to sighting error, and position tracking is required to compensate for it as a function of eye position in the cockpit [2]. Another potential benefit of precision position tracking of the helmet within the cockpit is

META-GNTX-00003970

coordinate intersection cueing, whereby the pilot is able to select or activate particular instruments within the cockpit by looking at them [3]. Modern applications involving enhanced or synthetic vision also require roll in order to present the horizon correctly, so any new HTS should be designed for full 6-DOF tracking capability, even though translation specifications are less critical and sometimes not given.

1.2. Current helmet-tracking technologies

After decades of evolution, magnetic trackers are the most mature technology, and have been used in many cockpit helmet-tracking programs. They have the advantages of a small head-mounted sensor and no line-of-sight requirement between source and sensor. One limitation is the very short range of magnetic fields. Since the dipole fields generated by the source fall off with the cube of distance, the resolution of the tracking system falls off with the fourth power of distance [4], limiting the high-performance tracking volume to a sphere of about 30 cm surrounding the source. This may not be a problem for tracking a single helmet in a small head-motion box, but sensitivity to distortion caused by metallic objects is very much an issue in the cockpit. Elaborate mapping procedures exist to compensate for these distortions, but they are time-consuming, and must be repeated after a change as small as adjusting the pilot's seat position. A tracking technology is strongly desired which will eliminate the need for cockpit mapping.

Various optical systems have been designed, typically using video cameras or electro-optical position sensing devices mounted in the cockpit viewing several infrared LEDs on the helmet. Such systems can be more accurate than the best magnetic trackers – and they don't require mapping. The main drawback is that optical sensors require a constant line-of-sight between the sensors and the LEDs, which can make it difficult to obtain full coverage and can lead to sudden data dropouts in the event of occlusion. It is also difficult to insure that the optical tracker will work flawlessly under the full range of lighting conditions from darkness to bright sunlight.

While latencies of both magnetic and optical tracking technologies have been improving, neither can provide high-quality prediction for more than a few milliseconds because of the jitter that results when numerically differentiating noisy signals.

To address these various problems, head-trackers have been developed based on miniaturized inertial sensors [5, 6]. Inertial sensors have no range limitations, no line-of-sight requirements, and no risk of interference from any magnetic, acoustic, optical or RF interference sources. They can be sampled as fast as desired, and provide relatively high-bandwidth motion measurement with negligible latency. Even tiny low-cost MEMS inertial sensors measure motion with very low noise resulting in jitter-free tracking which looks very smooth to the eye. Since they directly measure the motion derivatives, inertial trackers have the unique capability to perform high-quality prediction to compensate for graphics rendering and display rasterization latencies, without the noise problems of numerical differentiation.

Unfortunately, any slight inertial sensor bias or noise, when integrated over time, will cause the orientation and position outputs to gradually drift. With MEMS sensors suitable for helmet mounting, position drift is very fast, and even the orientation drift is problematic (tens or hundreds of degrees per hour). Thus, inertial sensors must be at least occasionally corrected by other sensors in a hybrid or aided-inertial tracking configuration. InterSense, has developed products which use tilt sensors and a compass to correct gyro drift in a sourceless orientation tracker (IS-300, introduced in 1996), followed by a hybrid tracking technology introduced in 1998 which fuses ultrasonic range measurements with inertial tracking for 6-DOF tracking applications. The newest version of this, called the IS-900 SimTracker, has rapidly become the most widely selected tracker for new simulator programs [7].

While the IS-900 SimTracker has proven the benefits of a drift-corrected inertial tracking solution for simulators, it is not directly suitable for use in cockpits. Firstly, the extremely high angular accuracy requirements call for an optical aiding technology instead of an acoustic one. Secondly, the inertial sensors measure motion relative to the ground, while the ultrasonic aiding sensors measure the position of the head-mounted microphones relative to the vehicle-mounted emitters. The Kalman filter expects both sets of measurements to be in the same coordinate frame, and would give erratic results if it were fed conflicting measurements.

In this paper we report on the first prototype of the FlightTracker, which is designed to overcome these limitations. In Section 2 we present a condensed summary of the differential inertial tracking algorithms, originally derived in [8], which have been implemented in the FlightTracker to allow it to compensate for platform motion. In Section 3 we describe the novel inside-outside-in optical system employed by the FlightTracker to obtain the extremely high accuracy and robustness needed for the cockpit tracking application. Section 4 describes the prototype system implementation, and Section 5 shows preliminary test results which strongly suggest that the target accuracy and robustness goals will be easily exceeded by the final system.

2. Differential Inertial Measurement

In this section, we develop the equations to compute the changes in orientation and position of a pilot's helmet relative to the aircraft by combining inertial data from a

META-GNTX-00003971

helmet-mounted tracking IMU and an aircraft-mounted reference IMU. This produces a differential inertial measurement of the head motion relative to the aircraft, which can then have its drift corrected by periodic optical measurements of the head position relative to the aircraft, using a very similar Kalman filter error estimator to that which was used in the IS-600 and IS-900 for tracking a head relative to the earth with a single inertial sensor corrected by earth-relative aiding sensors. We start by outlining ordinary inertial tracking relative to the earth to familiarize the reader with our notation and basic inertial tracking concepts, and then develop the more complex equations for tracking relative to a moving platform.

2.1. Inertial tracking relative to fixed platform

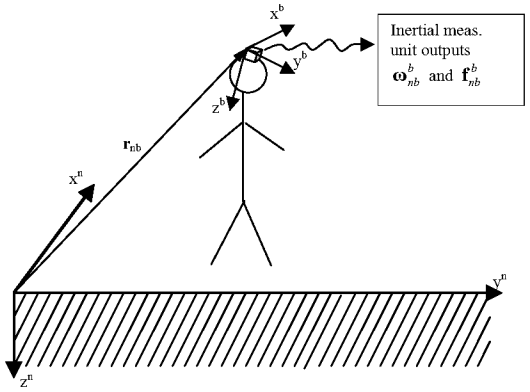

Figure 1: Inertial Tracking Relative to Stationary Ground

Figure 1 illustrates the simple case of using an inertial system to track the pose of a body, b, with respect to a fixed navigation frame, n. In this situation, which accurately represents the operation of current inertial tracking products, there are only two coordinate frames used. Vectors and matrices are designated with boldface characters, and superscripts, if present, indicate in which frame vectors are coordinatized. The subscripts on $\mathbf{r}_{nb}$ indicate that it is the position displacement vector from the n-frame origin to the b-frame origin. Likewise, $\boldsymbol{\omega}^b_{nb}$ represents the angular rate vector of the b-frame w.r.t. the n-frame coordinatized in the b-frame. This is exactly what the strapped-down triad of rate gyros aligned with the b-frame axes measures. The accelerometer triad at the b-frame origin senses $\mathbf{f}^b_{nb}$, the non-gravitational acceleration (a.k.a. specific force) of b-frame w.r.t. the inertial reference frame, n , expressed in b-frame.

The orientation of the b-frame w.r.t. the n-frame can be represented using a 3x3 rotation matrix $\mathbf{C}^n_b$ that transforms vectors from b-frame to n-frame: $\mathbf{v}^n = \mathbf{C}^n_b \mathbf{v}^b$. The orientation is integrated using the continuous-time (CT) differential equation:

$$\dot{\mathbf{C}}^n_b = \mathbf{C}^n_b S(\boldsymbol{\omega}^b_{nb}), \tag{1}$$

where $S(\boldsymbol{\omega}^b_{nb}) \equiv [\boldsymbol{\omega}^b_{nb}\times]$ is the skew-symmetric matrix formed from the elements of $\boldsymbol{\omega}^b_{nb}$ to implement the cross-product operator noted in the square brackets. The updated rotation matrix is then used to resolve the accelerometer readings into the n-frame, whence they can be easily corrected for the effect of gravity and double integrated to obtain the head position using:

$$\begin{aligned} \dot{\mathbf{v}}^n_{nb} &= \mathbf{C}^n_b \mathbf{f}^b_{nb} + \mathbf{g}^n \\ \dot{\mathbf{r}}^n_{nb} &= \mathbf{v}^n_{nb} \end{aligned} \tag{2}$$

where $\mathbf{g}^n \approx \begin{bmatrix} 0 & 0 & 9.8 m/s^2 \end{bmatrix}^T$ is the local apparent gravity vector which by definition points downward in n-frame.

Equations (1) and (2) are integrated numerically to keep track of orientation, velocity and position. They are simplified compared to those used in terrestrial inertial navigation, but the gyro sensors in the InertiaCube™ are not sensitive enough to detect the 15°/hr rotation of the earth, so there is no need to include terms to compensate for its effect on the sensors. The drift that results from using such low-performance gyros and neglecting the effects of earth rotation must be frequently corrected by other means, such as ultrasonic or optical measurements.

2.2. Inertial tracking relative to an arbitrary maneuvering platform

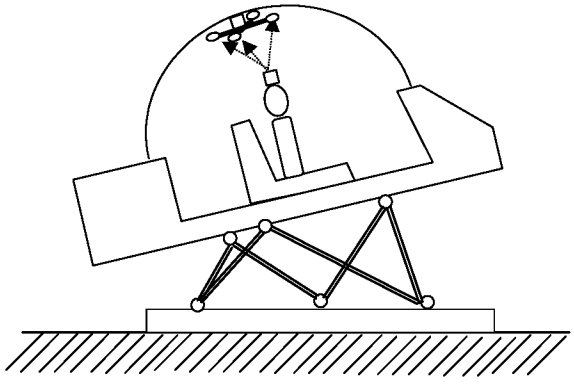

Figure 2: Tracking Head Relative to Platform

We will now attempt to solve the problem of tracking a person's head relative to a maneuvering platform by using "differential inertial measurements" between a first inertial measurement unit (IMU) mounted on the helmet and a second fixed on the aircraft frame. Figure 2 illustrates the hardware for such a system, installed in a motion-base simulator cab.

The first step to achieving the desirable tracking system illustrated in Figure 2 is to choose the n-frame fixed to the moving platform. We now have *three* coordinate frames to work with. The inertial frame, or i-frame, is any stationary inertial reference frame fixed w.r.t. the ground. As before, the b-frame is defined fixed

META-GNTX-00003972

in the body being tracked, normally the helmet, with the x-axis forward, y-axis right, and z-axis down. The n-frame is now a moving frame fixed relative to the aircraft. To simplify our equations, we will consider the n-frame to have its origin at the reference IMU and axes aligned with the reference IMU axes, as illustrated in Figure 3.

We wish to track the b-frame relative to the n-frame, which itself is moving relative to the inertial reference frame. The IMU on the helmet always measures the angular velocity and non-gravitational acceleration of the b-frame relative to i-frame, expressed in b-frame: $\boldsymbol{\omega}_{ib}^{b}$ and $\mathbf{f}_{ib}^{b}$. Likewise, the reference IMU on the aircraft always measures the angular velocity and non-gravitational acceleration of the n-frame relative to i-frame, expressed in n-frame: $\boldsymbol{\omega}_{in}^{n}$ and $\mathbf{f}_{in}^{n}$. Since the angular velocity vector relation $\boldsymbol{\omega}_{ib} = \boldsymbol{\omega}_{in} + \boldsymbol{\omega}_{nb}$ holds in any frame, we have in the b-frame:

$$\boldsymbol{\omega}_{nb}^{b} = \boldsymbol{\omega}_{ib}^{b} - \mathbf{C}_{n}^{b}\boldsymbol{\omega}_{in}^{n}, \tag{3}$$

which can be substituted into equation (1) and integrated to track the current orientation of the helmet w.r.t. the platform, $\mathbf{C}_{b}^{n}$. This may be solved directly, as $\boldsymbol{\omega}_{ib}^{b}$ is available from the tracking IMU, and $\boldsymbol{\omega}_{in}^{n}$ is measured by the reference IMU gyros on the moving platform.

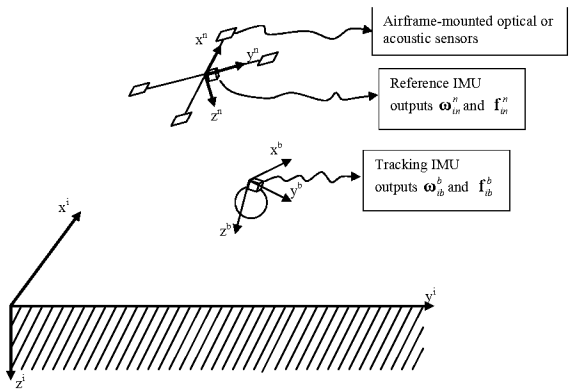

Figure 3: Coordinate Systems for Relative Inertial Tracking

To find the b-frame's velocity and position relative to the n-frame, which is rotating relative to the i-frame, we need to make use of the Law of Coriolis, which can be stated in its most general form as

$$D_a = D_b + \boldsymbol{\omega}_{ab}\times \tag{4}$$

where D_a represents an operator that differentiates any vector w.r.t. the a-frame, and a and b are any two Cartesian coordinate frames that share a common origin but are rotating with a relative angular velocity $\boldsymbol{\omega}_{ab}$. Start with the equations of motion in the inertial frame, which are exceptionally straightforward:

$$D_i^2\mathbf{r}_{ib} = \mathbf{f}_{ib} + \mathbf{g}_m \tag{5}$$

This is a vector relationship which will hold true in any coordinate frame, thus the lack of superscripts. $\mathbf{g}_m$ represents a pure mass attraction gravity vector. We now wish to convert this into n-frame. Expanding the left side of equation (5) with the Coriolis operator (4) we get:

$$\begin{aligned} D_i^2\mathbf{r}_{ib} &= D_i^2\mathbf{r}_{in} + \left\{D_i^2\mathbf{r}_{nb}\right\} \\ &= D_i^2\mathbf{r}_{in} + \left\{(D_n + \boldsymbol{\omega}_{in}\times)^2\mathbf{r}_{nb}\right\} \\ &= D_i^2\mathbf{r}_{in} + \left\{(D_n + \boldsymbol{\omega}_{in}\times)(D_n\mathbf{r}_{nb} + \boldsymbol{\omega}_{in}\times\mathbf{r}_{nb})\right\} \\ &= D_i^2\mathbf{r}_{in} + \left\{\begin{array}{l} D_n^2\mathbf{r}_{nb} + D_n\boldsymbol{\omega}_{in}\times\mathbf{r}_{nb} + \boldsymbol{\omega}_{in}\times D_n\mathbf{r}_{nb} \\ +\boldsymbol{\omega}_{in}\times D_n\mathbf{r}_{nb} + \boldsymbol{\omega}_{in}\times(\boldsymbol{\omega}_{in}\times\mathbf{r}_{nb}) \end{array}\right\} \\ &= D_i^2\mathbf{r}_{in} + \dot{\mathbf{v}}_{nb} + \dot{\boldsymbol{\omega}}_{in}\times\mathbf{r}_{nb} + 2\left(\boldsymbol{\omega}_{in}\times\mathbf{v}_{nb}\right) \\ &\quad +\boldsymbol{\omega}_{in}\times(\boldsymbol{\omega}_{in}\times\mathbf{r}_{nb}) \end{aligned} \tag{6}$$

Using equation (5) to substitute both $D_i^2\mathbf{r}_{ib} = \mathbf{f}_{ib} + \mathbf{g}_m$ and $D_i^2\mathbf{r}_{in} = \mathbf{f}_{in} + \mathbf{g}_m$ and rearranging terms, we get

$$\begin{aligned} \dot{\mathbf{v}}_{nb}^{n} &= \mathbf{C}_b^n\mathbf{f}_{ib}^{b} - \dot{\boldsymbol{\omega}}_{in}^{n}\times\mathbf{r}_{nb}^{n} - 2\left(\boldsymbol{\omega}_{in}^{n}\times\mathbf{v}_{nb}^{n}\right) \\ &\quad -\boldsymbol{\omega}_{in}^{n}\times\left(\boldsymbol{\omega}_{in}^{n}\times\mathbf{r}_{nb}^{n}\right) - \mathbf{f}_{in}^{n} \end{aligned} \tag{7}$$

where the orientation matrix $\mathbf{C}_{b}^{n}$ comes from integrating

$$\dot{\mathbf{C}}_b^n = \mathbf{C}_b^n S(\boldsymbol{\omega}_{ib}^{b} - \mathbf{C}_n^b\boldsymbol{\omega}_{in}^{n}) \tag{8}$$

and position $\mathbf{r}_{nb}^{n}$ is obtained by integrating

$$\dot{\mathbf{r}}_{nb}^{n} = \mathbf{v}_{nb}^{n}. \tag{9}$$

If the reference IMU is mounted at the origin of the n-frame, then it directly measures $\mathbf{f}_{in}^{n}$ and $\boldsymbol{\omega}_{in}^{n}$, so (7) is the complete navigation equation, which can be integrated using just data available from the two IMUs. (A numerical approximation to the derivative of $\boldsymbol{\omega}_{in}^{n}$ is needed if angular accelerometers are not available on the reference IMU). The second, third and fourth terms are new additions compared to the stationary platform equation (2). They represent tangential, Coriolis, and centripetal accelerations respectively which result from the rotation of the n-frame. Calculating the tangential acceleration term requires numerical differentiation of the discrete-time gyro outputs, and so could be a potential source of integration error if the lever arm $\mathbf{r}_{nb}^{n}$ from the reference IMU to the tracking IMU is too long.

By measuring the effect of gravity directly in the n-frame with a reference IMU, we do not need to know the platform tilt angles to resolve the gravity vector into the n-frame. Thus, this system operates without any inputs from the motion-base controller or the vehicle attitude reference system, and without the need to ever measure or calculate the pose of the moving platform.

3. Inside-Outside-In Optical Measurement

META-GNTX-00003973

3.1. Outside-in optical tracking

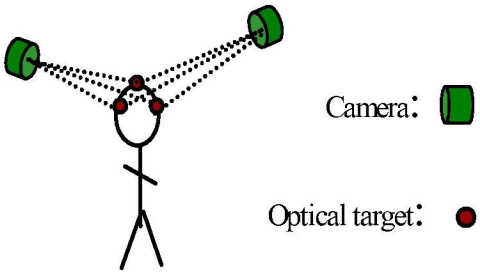

Figure 4: Outside-in optical tracking

Most commercial optical trackers, including the current generation of optical helmet trackers, are outside-in, which means the optical sensors are fixed around the outside of the tracking volume looking inwards at optical targets mounted on the tracked objects.

With 2-D sensors, a minimum of two outside-in sensors can be used to triangulate the position of a target, by essentially intersecting the rays from each camera towards the target. With 1-D optical sensors, a minimum of three sensors is needed to triangulate a target. Both methods are capable of triangulating the X,Y,Z location of a point target (such as an LED or a retro-reflective ball) to very high resolution and accuracy, typically better than 1 mm. However, for an outside-in tracker to obtain orientation, it must triangulate the positions of three point targets on the helmet, and solve for orientation from the positions. The quality of this orientation estimate is directly dependent on the quality of the position estimates and the separation between the point targets. With small separations between targets on a helmet, the orientation estimate is very sensitive to small errors in the position of any of the three targets. For example, with baseline separation of 100 mm, an error of 1 mm in one of the targets will cause an error of 10 mR in the orientation.

3.2. Inside-out Optical Tracking

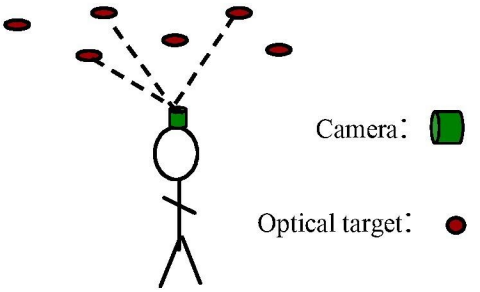

Figure 5: Inside-out optical tracking

In 1990, the University of North Carolina – Chapel Hill reported an optical ceiling tracker technology that can achieve milliradian angular accuracy [9]. This was accomplished by the use of "inside-out" optical tracking, in which the cameras are placed on the user's head and the targets fixed on the ceiling. This arrangement provides very high sensitivity to head rotations because even a small rotation will cause the LEDs on the ceiling to move substantially in the image plane of the cameras. Unfortunately, to also resolve position well requires use of multiple cameras aimed in different directions. The "HiBall" sensor cluster consists of a dodecahedral assembly of 6 lenses and 6 lateral-effect photo-diodes (LEPD), producing effectively 26 narrow fields of view [10]. This system was designed to provide high accuracy over a large area in a laboratory environment, but is not well suited for use in simulator environments or cockpits because 1) it requires a very dense array of bright IR LEDs as its targets, 2) reflective surfaces such as the canopy or screens in a simulator could cause significant errors if they create any reflection of an activated LED in one of the fields-of-view, and 3) the weight is prohibitive.

Considering the wisdom of the inside-out optical approach for applications requiring high angular precision and the limitations of the HiBall implementation, InterSense developed a new optical/inertial tracking technology to target mobile robot navigation and wearable augmented reality systems for large scale manufacturing, construction and aircraft maintenance [11]. The first prototype vision-inertial self-tracker (VIS-Tracker) was demonstrated at ISAR 2001. Since then, we have made many improvements in system architecture, hardware and software implementation. The VIS-Tracker system has now been demonstrated at numerous conferences and is now in beta-testing at several customer sites. To the best of our knowledge, the VIS-Tracker is the first hybrid optical/inertial tracking product.

3.3. Inside-outside-in Optical Tracking

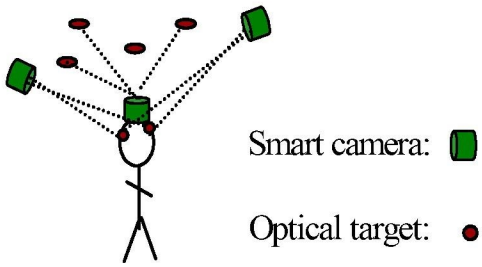

Figure 6: Inside-outside-in optical tracker.

Figure 4, Figure 5, and Figure 6 show three optical tracking configurations that can be created almost automatically using our flexible sensor fusion architecture [12]. Figure 5 shows an inside-out optical tracker using just one camera, like the VIS-Tracker. This configuration involves some significant performance trade-offs when choosing the field-of-view (FOV) of the lens. The geometric dilution of precision (GDOP) is a function describing the sensitivity of a tracking pose recovery

META-GNTX-00003974

algorithm (PRA) to small errors in the input measurements. The optimal GDOP in this configuration occurs when using three fiducials so widely separated that the lines of sight (doted lines in the figure) are all nearly orthogonal. To be able to select such widely separated fiducials in every frame requires using a very wide-angle lens, over 100° FOV. However, such a wide-angle lens has some disadvantages. The distortion in the edges becomes so extreme that it cannot be completely compensated. A measurement error of 0.2 pixels r.m.s. in determination of the centroid of a fiducial corresponds to a larger angular measurement error in mRads, since each pixel spans a larger angle. And the fiducials must be made larger to be readable at the same distance. On the other hand, a narrower lens can use smaller fiducials, but they need to be more densely spaced to make sure there are always several in view. Thus, practical lens choices for single-camera self-tracking range from about 60-90°. Since this is on the narrower side of the ideal GDOP, the position stability suffers, especially in the depth axis, while the angular stability is only affected a little bit, as a result of the lateral errors in position determination.

Figure 4 shows the classical outside-in configuration. Here the position of each marker on the head can be determined directly with sub-millimeter precision by triangulating the lines of sight from the two reference cameras. The orientation must be calculated using the position of three known markers on the head. Therefore, to achieve a high angular accuracy of 1 mR would require the system be designed to achieve positional accuracy of about 70 microns – much higher than actually required for position tracking purposes. To achieve this, the outside-in cameras will need narrow FOV lenses, which means the region of overlap or tracking volume will be small unless the tracking cameras can be set back a great distance from the tracking area, which isn't feasible in a cockpit.

The "inside-outside-in" configuration in Figure 6 offers excellent orientation and position tracking without requiring unduly high-resolution sensors, large stand-offs or large numbers of sensors. The outside-in sensors need just one head-mounted target visible to determine position to under 1 mm. Keeping one target visible for all head poses is much easier than guaranteeing visibility of three targets at all times, resulting in greater robustness with fewer targets or sensors. Orientation tracking is left to the single head-mounted sensor, which intrinsically offers extremely high angular resolution even with a modest resolution sensor. With the position already constrained by the outside sensors, there is no need to insure that the head-mounted sensor can see several widely spaced targets. In fact, it only needs to see one target to fully constrain the azimuth and elevation angles with a precision equal to the angular resolution of the camera, a fraction of a milliradian. The roll, which is less critical for display registration, will be determined to sufficient accuracy by gravity, and to even higher accuracy if the head mounted camera can see two targets in the cockpit or if the outside-in cameras can see two targets on the helmet. Therefore, one can select almost any FOV for the inside-out camera and still obtain outstanding performance in all 6 degrees of freedom. The drawing shows two outside-in cameras, as this is enough to completely constrain the head position without any use of the inside-out camera. In concert with the inside-out camera, even one outside-in camera would be sufficient to lock up all DOFs except for translation along the outside camera's axis. Thus two outside cameras is already highly redundant, allowing great robustness for occlusions.

High position and orientation accuracy can also be obtained by using a large number of head-mounted cameras pointing in many directions (like the HiBall) and a large number of beacons all around the space so that the cameras can all find targets. However, there are two large advantages of the "inside-outside-in" strategy. First, it involves much less head-mounted weight – just a single camera and lens, which can ultimately be a tiny CMOS sensor module of the type now being produced for cell phones. Second, there is far less infrastructure to install. The head-mounted camera just needs to find one target somewhere within its, say, 90 deg. FOV. Just four LED targets would allow a full +/- 180 deg. range of head yaw. This is a very desirable property for a system to be installed in tightly packed cockpits where mounting spaces are scarce.

Hoff originated the idea of fusing data from inside-out and outside-in optical sensors in [13]. In that work, the pros and cons of inside-out and outside-in optical systems were identified, and the benefit of combining pose from an outside-in and an inside-out tracker was recognized. Elegant equations were derived for combining two pose estimates, weighting the estimates according to the covariance matrices describing the relative uncertainty in each estimate. A simulation was provided showing the improvement in an AR scenario in which pose data from an outside-in optical tracker (Northern Digital Optotrak) was combined with pose data obtained by solving a pose-recovery algorithm from an inside-out camera viewing 5 LED targets. This method works by first computing the 6-DOF pose from the outside-in tracker and the 6-DOF pose from the inside-out tracker and then fusing the two 6-DOF pose estimates. While this suffices to demonstrate the accuracy improvement that can be obtained by the inside-outside-in method, it does not reap the full benefit in robustness that can be obtained, since both trackers must see enough targets to provide 6-DOF pose estimates. In our implementation, the system will automatically fuse any 2-D bearing measurements obtained from any of the outside-in cameras and any inside-out 2-D bearing measurements with the pose being maintained by the

META-GNTX-00003975

inertial sensors, using a real-time Extended Kalman Filter. Therefore, it is not necessary for the inside-out camera to see a complete set of targets in order to provide useful data, nor is it necessary for any head-mounted target to be seen by more than one outside-in camera in order for its measurement to be used. We can block all the inside-out targets, or all the outside-in targets, or most of both sets of targets, and only slightly affect the tracking quality.

Satoh et al [14] recently presented an implementation of the inside-outside-in concept in which they fused data from a single outside-in camera and a single inside-out camera. This is a specific instance of the flexible system described below, but it does illustrate the improvement in robustness as well as accuracy. We thoroughly agree with Satoh et al's analysis and find that their results convincingly reinforce Hoff's thesis that combining inside-out and outside-in tracking is highly advantageous. With due respect, we claim that our implementation is more flexible, faster, and more accurate.

4. Prototype Implementation

Figure 7 shows an overview of the system that was used to collect the results in the next section. The key hardware component is the InertiaCam™ device originally developed for the VIS-Tracker product. The InertiaCam consists of a miniature sensor head and an image processing unit (IPU). The sensor head (26 x 15 x 49 mm) contains the inertial sensors, CCD and lens. The IPU contains a low-power DSP and interface electronics.

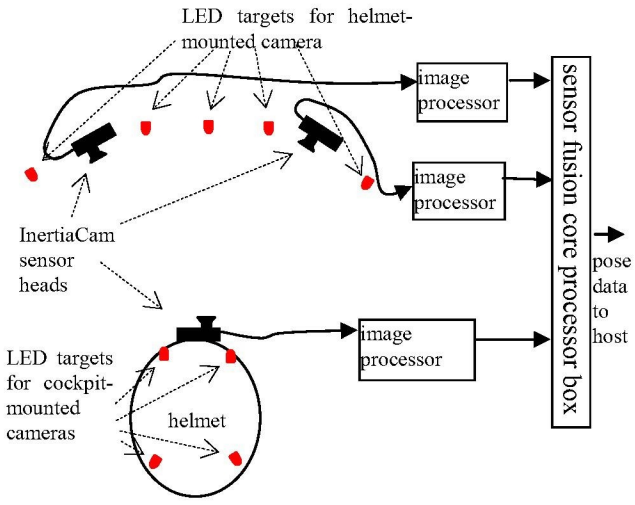

Figure 7: Schematic overview of tested system configuration

For the VIS-Tracker, the DSP was programmed to automatically detect and decode circular fiducial patterns. Fiducial patterns or even natural features in the environment are desirable optical targets for a wearable self-tracker which needs to operate over an area too large to install active targets throughout. However, in the cockpit, active targets such as LEDs are practical, take up less space then patterns, and offer more robust detection under cockpit lighting conditions. We have modified the firmware in the InertiaCam to decode and track LEDs instead of fiducials. To automatically acquire initial pose, the system needs to identify individual LED targets and look up their pre-stored positions. We have invented a novel scheme of modulating the LEDs with ID codes, which requires no synchronization between the LEDs or the cameras, and can be decoded simply in the InertiaCam firmware during acquisition.

The Sensor Fusion Core (SFC) firmware based on the architecture of [12] was implemented using the Mathworks' Real-Time Workshop with a custom-developed target for the ETS real-time operating system, running in an embedded x86 processor box. During initialization, the SFC invokes the meta-driver to enumerate all the sensor and target hardware attached to the system, and then reads configuration files (or solicits user input if necessary) to determine which cameras are used inside-out and which are used outside-in. The system will support an arbitrary number of cameras in each role, although if there are less than two outside-in cameras, there must be at least one inside-out.

Having established a basic configuration structure, the SFC architecture is designed to auto-calibrate the exact poses of all sensors and targets. We have previously demonstrated position auto-mapping, but have not yet implemented full 6-DOF auto-calibration as would be required to get the extrinsic parameters of the cameras in the cockpit. Therefore, for this evaluation we mounted the outside-in cameras on a precision machined aluminum bar to hold them at exactly known positions and angles, and manually measured the LED positions relative to the InertiaCam sensor package on the helmet assembly.

During operation in the aircraft, the system starts out in acquisition mode, and then switches to tracking mode as soon as it is able to acquire an initial pose estimate. In the unlikely event that the tracking gets lost (e.g. the helmet is removed from the tracking area for an extended time), the system will automatically return to acquisition mode and re-acquire, typically in less than a second. During acquisition, the SFC scheduler requests each sensor to perform a "scan" and report all targets it can see and decode. The sensor/target measurement results are stacked in an accumulator, and if there is a sufficient combination the system solves for pose, otherwise it re-starts the scan. Minimum sufficient combinations are either 4 inside-out measurements (not all 4 targets collinear), 4 outside-in (at least 3 targets, not all collinear), 1 outside-in plus 2 inside-out, or 3 outside-in plus 1 inside-out.

During tracking mode, the system performs one scheduling operation, one inertial update and Kalman propagation step, and one Kalman measurement update (if there is a new measurement return available) on each

META-GNTX-00003976

cycle, running at 180 cycles/second. Because of the total generality of the SFC architecture, the scheduler is actually the most complex and computationally intensive part. For each sensor, the system finds all potential targets which are currently within its field-of-view and oriented so as to be usable. The list of all potential sensor-target pairs is submitted to an arbitration server to coordinate use of shared sensor/target resources between multiple SFCs on the same system or operating in the same area. (In the simple example here with only one SFC, all the potential sensor-target pairs will be marked usable by the arbiter). Next, for each usable i-j pair, the scheduler computes an expected benefit value, B_{ij}, and the pair with the highest value gets triggered to make a measurement. Ideally, the expected benefit calculation would involve computing the Kalman filter information gain or covariance reduction as a result of incorporating the contemplated measurement, but that would be far too costly to compute for each of possibly hundreds of potential sensor-target pairs. Therefore, we developed some heuristic algorithms to approximate the benefit based on such things as distance from recently used targets, measurement noise, and likelihood of completion.

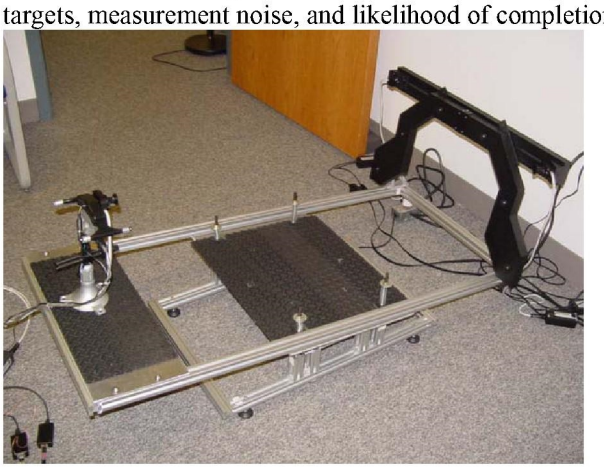

Figure 8: InertiaCams and LEDs on cockpit motion testbed

Figure 8 shows the 2 outside-in InertiaCams, and 6 LED targets mounted on a "canopy bow" on the back of a motion platform testbed, with a helmet-tracking assembly consisting of an inside-out InertiaCam and 4 LED targets attached on top of a firewire camera at the front of the platform. The platform is driven by a motorized cam that generates cyclical motion in pitch and roll, used to test the efficacy of the differential inertial tracking algorithms.

5. Evaluation

5.1. Video augmented reality registration demo

We developed an augmented reality (AR) program to demonstrate the registration accuracy of the tracking system. This program receives video frames from a firewire camera, applies them as a background texture map in OpenGL, and then renders foreground objects according to the camera viewpoint as reported by the motion tracking system. The geometry of the foreground "augmentations" are determined by creating a 3D model in 3DStudio and exporting a file in .3ds format. This file is loaded by the AR demo program and rendered at 60 frames per second from the dynamically changing viewpoint received from the tracker. The result is a window showing the video from the firewire camera, augmented with virtual objects which appear to be superimposed on top of corresponding real objects. For the demo, we created a 3D model of the "canopy bow", with a photograph of cockpit instrumentation added in the interior space, which was superimposed on top of the real black metal "canopy bow". The registration appeared very good, on the order of 1-3 mm apparent displacement. There was almost no perceptible lag or dynamic mis-registration. We recorded video of the AR overlay demo, which gives a good subjective impression of the registration accuracy both static and dynamic. However, we have yet to find a good way to quantify this registration accuracy. From the video, we know that the errors appear to be unbiased- the overlay is sometimes shifted a little left, sometimes a little right, but on average it is right on top of the real object. For now we are using just stability, jitter and repeatability measurements to characterize the tracking system.

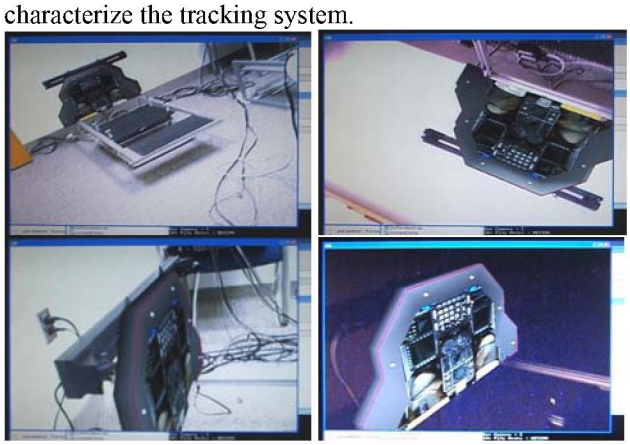

Figure 9: Stills from AR demo video

Figure 9 shows some stills from the video illustrating the superimposition of the graphical model of the "canopy bow" on top of the background video coming from the firewire camera. The second shot shows a view when the helmet tracking sensor has been rotated upside down. The third shows a view from the extreme side (basically outside the left window of the cockpit) where the helmet assembly is already outside the view of both outside-in cameras, and tracking is working off the inside-out camera alone. The fourth view is taken with the lights off. A short clip from the video has been submitted and posted at www.isense.com/support/downloads/FlightTrackerDemo.zip. We strongly encourage you to view it because stills

META-GNTX-00003977

cannot capture the dynamic registration effect, which is the most impressive part of the demo.

5.2. Stability tests

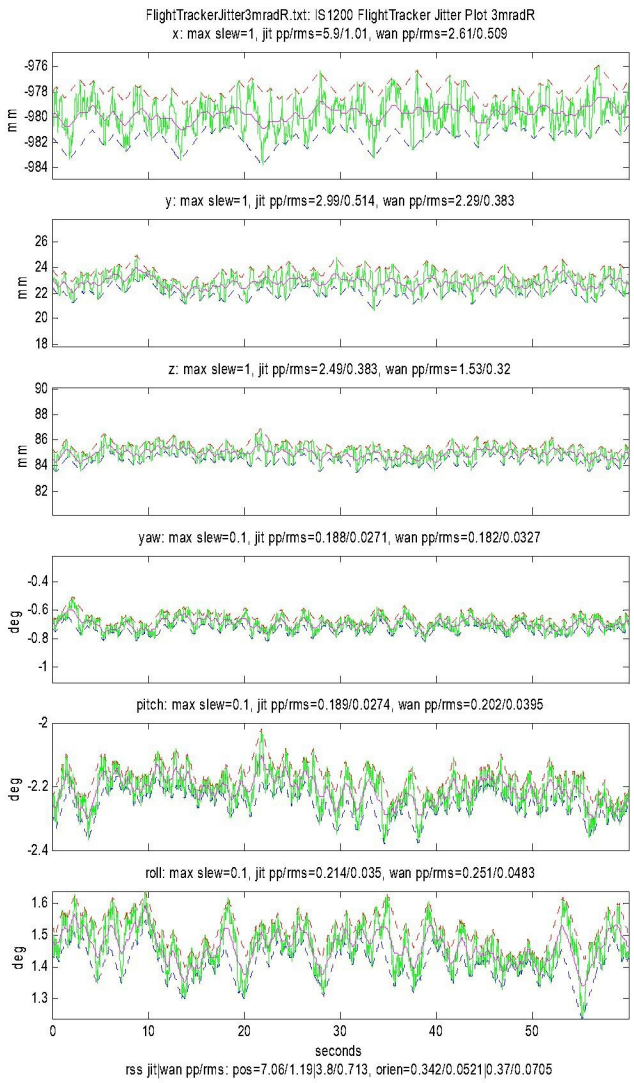

Figure 10: Jitter and wander of tracker outputs at rest

Figure 10 shows the recorded output data from the tracker for approximately one minute while the tracker sensor was resting on a Panavise on the motionless platform, at a fixed location approximately one meter back from the bar with the two outside-in cameras. To save space we omit the plot of a minute of data recorded a few minutes later, under the same conditions but with the platform undergoing simulated aircraft motion driven by the motor. The second plot exhibits slightly higher jitter and wander, but both experiments measured the same average values for pose of the head-tracker relative to the bar. The moving platform case exhibits higher wander especially in roll, which is the axis of maximum angular acceleration for the motion platform. The actual data is green. The red and blue bounds show upper and lower peak bounds used for measuring the high frequency "jitter". The magenta trace represents the "wander", defined as any drift slower than 1 mm/s or 1deg/minute.

Table 1 summarizes the results from these two experiments. The values in mm represent root-sum-squares of the x, y, and z components, and the values in degrees are root-sum-squares of yaw, pitch and roll components. Note that the r.m.s. stability is already very close to our target accuracy specs of 1 mm and 0.06 degrees. We therefore expect to easily exceed our ambitious accuracy goals in the final product. More work will need to be done to evaluate the absolute accuracy, the dynamic accuracy, and to quantitatively measure the registration accuracy.

Table 1: Stability results on fixed v. moving platform

	Fixed platform	Moving platform
r.m.s. jitter	1.19 mm 0.05 deg	1.39 mm 0.07 deg
Peak jitter	3.53 mm 0.17 deg	4.49 mm 0.23 deg
r.m.s. wander	0.713 mm 0.07 deg	1.01 mm 0.10 deg
Peak wander	1.9 mm 0.18 deg	2.67 mm 0.24 deg

Table 1 illustrates that the differential inertial tracking concept basically works. The errors increase slightly when the platform moves, but not more than expected. Without inertial sensors on the platform to cancel out the platform motion, the tracking would totally fail. By estimating platform angular acceleration by combining linear accelerometer data from both reference IMUs, we expect to further reduce the errors due to platform motion.

The table does not convey the robustness of the system. We were able to shake the head tracker around vigorously, turn it upside-down, block the inside-out camera completely and indefinitely, block BOTH outside-in cameras indefinitely, block all three cameras for a few seconds, and withstand all kinds of optical clutter in the fields of view of all the cameras, all without losing track. We can turn the head-tracker around completely, and the tracking system automatically transitions to using the two LEDs on the back of the helmet and continues to track. The helmet can be manipulated through any orientation (360 degrees in every rotation axis), and can be translated from within inches of the reference bar to about 8 feet away. It can be moved far off to either side where it is out of view of one and then both of the reference cameras, and will continue to track if it is facing forward. This hybrid tracker exhibits much greater robustness than pure

META-GNTX-00003978

optical trackers and much greater accuracy than magnetic ones (with no need for mapping).

6. References

[1] Mulholland, F. (2002). Helmet-Mounted Display Accuracy in the Aircraft Cockpit. In Proceedings of SPIE Helmet and Head-Mounted Displays VII, vol. 4711, Orlando, FL.

[2] Task,H.L. & Goodyear, C. (1998). Effects of Aircraft Windscreens and Canopies on HMT/D Aiming Accuracy: Part 3. In Proceedings of SPIE Helmet and Head-Mounted Displays IV, vol. 3689, Orlando, FL.

[3] Task, H.L. & Kocian,D.F.(1995). Design and Integration Issues of Visually-Coupled Systems. Wright-Patterson AFB. Armstrong Laboratory technical report AL/CF-SR-1995-0004.

[4] Nixon,M.A. et al (1998). The Effects of Metals and Interfering Fields on Electromagnetic Trackers. Presence, vol. 7, issue 2, pp. 204-218

[5] Foxlin, E. (1993). "Inertial Head-Tracking", M.S. Thesis, Dept. of EECS, MIT.

[6] Foxlin, E. Inertial Orientation Tracker Apparatus Having Automatic Drift Compensation for Tracking Human Head and Other Similarly Sized Body. US Patent 5,645,077, July 8, 1997.

[7] Delaney, Ben (2002). The Market for Visual Simulation & Virtual Reality Systems, Fifth Edition. ISBN: 1-929696-06-X. CyberEdge Information Services, New York, NY. www.cyberedge.com

[8] Foxlin, E. (2000). Head-Tracking Relative to a Moving Vehicle or Simulator Platform Using Differential Inertial Sensors. Proceedings of Helmet and Head-Mounted Displays V, SPIE Vol. 4021, Aerosense Symposium, Orlando, FL.

[9] Wang, J.F., Chi, V., and Fuchs, H. (1990). A Real-Time Optical 3D Tracker for Head-Mounted Display Systems. Symposium on Interactive 3D Graphics, Snowbird, UT.

[10] Welch, G., Bishop, G., Vicci, L., Brumback, S., Keller, K. & Colluci, D. (2001). High-Performance Wide-Area Optical Tracking: The HiBall Tracking System. Presence: Teleoperators and Virtual Environments vol 10, issue 1, MIT Press.

[11] Foxlin, E. & Naimark, L. (2003). VIS-Tracker: A Wearable Vision-Inertial Self-Tracker. IEEE Conference on Virtual Reality (VR 2003). Los Angeles, CA.

[12] Foxlin, E. (2002). Generalized Architecture for Simultaneous Localization, Auto-Calibration, and Map-building. IEEE/RSJ Int. Conf. Intelligent Robots & Systems, Lausanne, Switzerland.

[13] Hoff, W. (1998). Fusion of Data from Head-Mounted and Fixed Sensors. Proceedings of IWAR '98, San Francisco.

[14] Satoh, K., Uchiyama, S., Yamamoto, H. & Tamura, H. (2003). Robust Vision-Based Registration Utilizing Bird's-Eye View With User View. Proceedings of ISMAR 2003, Tokyo.

META-GNTX-00003979